WHAT IS CLAIMED IS:

1. A method of searching a document having nested-structure document-specific markup, the method comprising:

receiving a query that designates at least (A) a phrase to be matched in a phrase matching process, and (B) a selective designation of at least a tag or annotation that is to be ignored during the phrase matching process;

deriving query-specific indices based on query-independent indices that were created specific to each document; and

carrying out the phrase matching process using the query-specific indices on the document having the nested-structure document-specific markup.

2. The method of Claim 1, wherein the query-independent indices were created by a method including:

a) labeling elements in the document with intervals, wherein:

a1) for markup tags, the intervals are defined in terms of a starting index number associated with an opening markup tag and an ending index number associated with a closing markup tag that corresponds to the opening markup tag, and

a2) for single words, the intervals are defined in terms of a single index number associated with the word; and

b) forming the query-independent indices so that they are configured to be used in the searching method by first receiving, for a word or tag in the document, a position in the document, and by then indicating whether or not the word or tag is present at that position.

3. The method of Claim 2, wherein the step of deriving the query-specific indices involves deriving the query-specific indices from the query-independent indices without rebuilding any of the query-independent indices.

4. The method of Claim 3, wherein the step of deriving the query-specific indices includes forming at least one of a group including:

an index of each word in the phrase to be matched by the phrase matching process;

an index of context tags that may be found in the document; and

an index of at least a tag or annotation to be ignored during the phrase matching process.

5. The method of Claim 2, wherein the phrase matching process includes:

for each context interval, defined by a beginning index defining a position of beginning tag and a closing index defining a position of a closing tag, performing an index-nested loop by probing an index of each phrase word in order, and an index of each tag or annotation to be ignored, so as to construct at least one witness;

wherein each witness is a contiguous sequence of intervals contained within the context interval and includes each phrase word occurrence exactly once and in phrase order.

6. The method of Claim 5, wherein at least one witness includes each phrase word occurrence exactly once and in phrase order, interleaved with tags or annotations to be ignored.

7. The method of Claim 2, wherein the phrase matching process includes:

scanning, in document order, a combined index of (A) phrase words and (B) tags or annotations to be ignored, while using a stack to keep track of nested context intervals and annotation intervals;

wherein:

the stack includes at least one entry corresponding to a current context interval in which witnesses are identified; and

the at least one entry maintains a set of (A) partial witnesses that are being identified and (B) complete witnesses that have been identified, within the current context interval.

8. The method of Claim 1, wherein the query further designates:

a set of context tags defining a context to which the phrase match should be restricted.

9. The method of Claim 1, wherein:

the document's nested-structure document-specific markup is in Extensible Markup Language (XML).

10. The method of Claim 1, wherein:

the receiving step includes receiving a query that designates at least a phrase to be proximity-matched in the phrase matching process; and

the phrase matching process involves proximity phrase matching as distinguished from exact phrase matching.

11. A method of creating query-independent indices suitable for use in searching a document having nested-structure document-specific markup, the method comprising:

a) labeling elements in the document with intervals, wherein:

a1) for markup tags, the intervals are defined in terms of a starting index number associated with an opening markup tag and an ending index number associated with a closing markup tag that corresponds to the opening markup tag, and

a2) for single words, the intervals are defined in terms of a single index number associated with the word; and

b) forming the query-independent indices so that they are configured to be used in the searching method by first receiving, for a word or tag in the document, a position in

the document, and by then indicating whether or not the word or tag is present at that position.

12. The method of Claim 11, wherein:

the document's nested-structure document-specific markup is in Extensible Markup Language (XML).

13. A computer program product including computer executable code or computer executable instructions that, when executed, causes a computer to perform the governing step of Claim 1.

14. A computer program product including computer executable code or computer executable instructions that, when executed, causes a computer to perform the governing step of Claim 5.

15. A computer program product including computer executable code or computer executable instructions that, when executed, causes a computer to perform the governing step of Claim 7.

16. A computer program product including computer executable code or computer executable instructions that, when executed, causes a computer to perform the governing step of Claim 11.

17. A system configured to perform the method of Claim 1.

18. A system configured to perform the method of Claim 5.

19. A system configured to perform the method of Claim 7.

20. A system configured to perform the method of Claim 11.